DOCUMENT RESUME

ED 186 854 CS 005 433

AUTHOR Rakes, Thomas A.; McWilliams, Lana J.
TITLE A Comparison of Performance on Cloze Tests, Group Reading Inventories and Standardized Reading Achievement in Grades Seven, Eight and Nine.
PUB DATE Nov 79
NOTE 22p.; Paper presented at the Annual Meeting of the College Reading Association (23rd, Boston, MA, November 1-3, 1979).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cloze Procedure; Comparative Analysis; *Content Area Reading; *Informal Reading Inventories; Junior High Schools; Reading Ability; *Reading Achievement; *Reading Diagnosis; Reading Research; Reading Skills; Social Studies; Standardized Tests

ABSTRACT

A random sample of 300 seventh, eighth, and ninth grade students participated in a comparative study of performance on social studies cloze tests, social studies group reading inventories (GRI), and a popular standardized test (Gates-MacGinitie Reading Tests, Survey E). It was found that cloze tests and GRIs constructed from social studies content differentiated instructional/independent level readers from frustration level readers. The cloze tests and GRIs appeared to measure different reading skills. The data also indicated that readers' performance on cloze tests and GRIs was significantly different from performance on a group standardized test, such as the Gates-MacGinitie test. It was concluded that teachers concerned with student textbook reading ability might try using informal types of reading assessment in addition to standardized reading achievement tests. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED186854

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A Comparison of Performance on Cloze Tests, Group Reading Inventories and Standardized Reading Achievement In Grades Seven, Eight and Nine*

Thomas A. Rakes

Lana J. McWilliams

Learning Center
Department of Curriculum
& Instruction
College of Education
Memphis State University
Memphis, Tennessee 38152

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Thomas A. Rakes

Lana J. McWilliams

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS005433

*Abridged Report: July, 1979: Funding was provided in part through Memphis State University and The State of Tennessee.

A Comparison of Performance on Cloze Tests, Group Reading Inventories and Standardized Reading Achievement In Grades Seven, Eight and Nine

Introduction

Minimum competency testing in education has become a nation-wide movement. As a result of public opinion and legislative action, many states have taken steps to either in deed, or by decree, endorse proficiency testing. Recognizing significant trends in education, commercial publishers are producing various types of assessment batteries to be used in testing programs. In some states these tests have been developed by state-level committees.

In their traditional format, formal assessment batteries provide only overall scores, usually stated in such broad terms as "vocabulary" or "comprehension," with no pertinent or in-depth information regarding the instructional adjustments and considerations that need to be made. Several other problems are evident when using standardized test data for instructional purposes in classrooms.

1. An artificial relationship exists between test items and actual material found in adopted textbooks.
2. Many students, particularly those experiencing reading and/or learning problems, have experienced frustration while taking standardized tests. Their subsequent academic abilities.
3. Standardized test batteries are expensive when compared to informal procedures.

4. Most norm-referenced group standardized tests are more survey than diagnostic in nature. Even newer criterion referenced tests either sample a broad range of skills on a superficial basis or measure a select number of skills and exclude others on a primarily arbitrary basis. Both types usually use a variety of content related reading matter with insufficient depth to be of instructional value to most classroom teachers. Few classroom teachers will support such instruments as appropriate or particularly valuable in developing their teaching strategies.
5. Most standardized tests require longer test intervals than informal procedures.

The Need for Appropriate Diagnostic Instruments

The range of student reading achievement can span eight or more grade levels in a given middle school (Allington, 1975) and high school reading achievement may vary even more than eight years. This difference/deficit creates an almost unmanageable situation for content teachers. Above grade four, silent reading tasks are required and, often represent the single most necessary skill area needed for successful completion of a high school education. Empirical evidence exists to indicate that student reading achievement, teacher judgments of readability and publisher readability estimates of content textbooks may vary as much as six grade levels within a given text or study material (Berger, 1947; Cox, 1970; Galloway, 1973;

Rakes, 1973; Ramsey, 1961; Smith, Guire and Cheek, 1972; and Wood, 1954). Thus, an accurate measure of reader-text appropriateness is essential before instructional procedures can be developed.

Much has been written regarding several possible methods or instruments for surveying reading skills. However, few suggested procedures specifically identify reading-content related skills but, instead, screen for general study or basic reading skills (vocabulary, comprehension, and total reading) and the resulting grade or accuracy level score provides little information relative to specific instructional needs. Techniques most often suggested for use in grades 4-12 include the following: (1) cloze tests, (2) standardized tests, (3) informal reading inventories, and (4) group reading inventories. Two of these, cloze tests and group reading inventories, represent viable alternatives for use with groups of students.

Cloze Testing

Using a cloze technique (Bormuth, 1968) requires the selection of a passage (minimum of 250 words in length) from a context textbook. The representative selection is then retyped deleting every nth word excluding the first and last sentences. Students are then asked to read the cloze passage and fill in the blanks. Much research has been reported using a cloze procedure, most of which reflects a high degree of support for the procedure as a quick placement technique. The resulting information indicates to a teacher whether a specific book is written on the appropriate level for a given reader - much more relevant information than a general standardized score for a teacher who must use that book in teaching his/her content material.

Although this informal testing technique has been widely accepted and endorsed among reading specialists, the correlation of content area cloze tests scores with standardized reading scores has not been sufficiently established.

Group Reading Inventories

In addition to the cloze procedure, some reading consultants also recommend another type of testing, generally known as a Group Reading Inventory (GRI), or Content Inventory (CI), as an appropriate means of evaluating content related reading ability. A review of published literature reveals little empirical study (Rakes and McWilliams, 1979). What is available represents descriptive type summaries with sample GRI's (Burron and Claybaugh, 1974; Carvell, 1980; Dishner and Readence, 1977; Kaiser, 1975; Hafner, 1977; Koenke, 1972; Marksheffel, 1966; McWilliams and Rakes, 1979; Miller, 1974; Rakes, 1975; Rakes and McWilliams, 1978; Shepherd, 1973; Strang, 1964; and Viox, 1968). Not only is the GRI literature limited but it is characterized by a lack of information concerning the validity of GRI's. The empirical basis to support the use of Group Reading Inventories as a useful group screening technique, as suggested by several authors, is absent.

A Group Reading Inventory is an informal test administered to determine if specific reading materials are appropriate, too easy or too difficult for a given group of readers. The test is given by requiring a group of readers to read silently a 400-500 word sample that has been grade-leveled and selected for its content relation to a particular class. The readers are then asked to answer various comprehension level questions (i.e. main

ideas, context, details, and inference), without looking back or reviewing the material. The resulting score not only reflects the general reading level of each student but also indicates what particular comprehension tasks are most difficult for individual readers.

Ideally, group reading inventories and cloze tests should be developed from passages taken from the textbooks intended for use in subject area classrooms. The same course (classroom materials) has, for many years, been recommended for the "proper" use of informal reading inventories. Experience has shown however, that teachers and many reading specialists prefer to use commercially prepared IRIs. Therefore, the only commercially available, content area group reading inventories and cloze tests were selected as representative of such informal tests. Content Inventories: English, Science and Social Studies (McWilliams and Rakes, 1979) is the pioneer effort for such group, content area tests and seemed appropriate for use in this study.

Although group reading inventories are recommended by reading specialists as an alternative to standardized testing, until recently no prepared inventories were available for content teachers. In addition, literature referring to or concerned with a GRI reveals that empirical support for the use of a GRI is relatively limited.

To date, there exists little evidence to verify that these types of tests using subject area content or even this concept of testing is valid as an alternative method of assessment.

Purpose of the Study

The following questions were considered:

1. Does a significant difference or correlation exist between reader performance on social studies cloze tests and group reading inventories?
2. Does a significant difference or correlation exist between reader performance on a standardized test and performance on social studies cloze tests and group reading inventories?

Methodology

Subjects

A random sample of 300 subjects was taken from one public and one private school located in an urban area in the southern portion of the United States. A total of 85 seventh, 86 eighth, and 129 ninth grade subjects were selected for participation in the study. A total of 286 subjects complete all tests. Mean comprehension and vocabulary grade scores on the Gates-MacGinitie Reading Test were 9.3 and 8.9 (seventh grade subjects), 10.2 and 9.9 (eighth grade subjects) and 10.7 and 10.6 (ninth grade subjects).

Instruments

The Gates-MacGinitie Reading Tests, Survey E (Gates and MacGinitie, 1965) was administered to all subjects. The standardized test measures reading performance in the areas of Speed, Accuracy, Vocabulary and Comprehension. Cloze tests and group reading inventories were selected from Content Inventories: English, Social Studies and Science (McWilliams and Rakes, 1979).

Using a 7th word deletion pattern the following three social studies cloze tests were administered:

"Democracy" - grade 7 (50 deletions)

"Formation of Political Parties in a New Nation" - grade 8 (51 deletions)

"Labor Unions" - grade 9 (50 deletions)

Three social studies group reading inventories were administered:

"Russia" - grade 7 (410 total words)

"The Gold Rush" - grade 8 (497 total words)

"Money" - grade 9 (519 total words)

Each GRI includes fourteen questions including 2 main idea, 3 context, 5 detail and 4 inference type questions. All tests were administered on an alternate basis by rotating the order of administration of each of the three measures. Procedures used for administration and interpretation of the tests were as recommended by the authors of the tests.

For the cloze tests, percentage correct scores were used to determine instructional reading levels: 36% or lower - frustration level, 37% - 57% - instructional level, and 58% or higher - independent level. Placement levels for reader performance on a group reading inventory containing fourteen questions were determined by using the following range: 86% or higher (12-14 questions correct) - independent level, 79% - 64% (9-11 questions correct) - instructional level, and 57% or lower (8 or fewer questions correct) - frustration level. The criteria for both cloze and GRI placement levels were suggested by McWilliams and Rakes (1979).

Results

A Pearson product-moment correlation coefficient was used as a measure of association indicating strength of relationship of selected variables. Data were compared on a grade by grade and total sample basis. Within each grade comparisons were made between cloze test performance, group reading inventory performance, and the raw test scores of the three subtests and one total score on the Gates-MacGinitie Reading Tests, Survey E (GMRT). The relationship between the types of questions missed on a social studies GRI (main idea, detail, context and inference) was also compared to cloze test and the GMRT performance. The .001 level of significance was used as the criterion for all correlational comparisons. Most correlations were found to be statistically significant. However, many correlations, even though statistically significant, were not sufficiently strong enough to indicate the existence of a dependable relationship.

For the total sample of subjects in grades 7,8, and 9, strong correlations were found between the types of questions missed and GRI percent correct, number correct and placement levels. Lower correlations were found between GRI questions and cloze test placement levels. The single exception was found between missed context questions and both cloze and GRI placement levels. For this comparison .45 and .34 correlations were found for GRI and cloze percent correct scores. See Table I.

When the same information was compared with the subtest raw scores of the GMRT, correlations of .48 or above were found for the comprehension

sub-test scores. Cloze test placement levels were also statistically higher and on vocabulary and comprehension reached a level of .53 and .48 respectively. The lowest correlations throughout the analysis occurred between GRI context questions missed and the other variables.

Similar results were found on a grade by grade analysis. Most correlations were significant but many fell below an r value of .44. The GMRT comprehension subtest scores correlated with cloze and GRI placement levels at a level higher than the other three GMRT sub-test scores. See Table II.

A one-way analysis of variance was used to test for a significant difference between cloze test placement levels and GRI placement test levels for each of the GMRT sub-test scores. On a grade by grade basis, results are provided in Tables III-VIII. In an attempt to maintain the exact probability, .05 was reduced to .0125 (.05/4) to further reduce the probability of spurious significance.

For the seventh and eighth graders a significant difference was indicated for both cloze and GRI placement levels for vocabulary and comprehension sub-tests. The ninth grade sample resulted in a significant difference between GRI and cloze placement levels for all four sub-test areas on the GMRT.

Discussion

The correlation data indicates that although statistically significant r values exist for nearly all comparisons, many fell within the .44 to .55 range. The high correlations between three of the four types of questions used on the group reading inventory (main idea, details and inference)

and the GRI placement level indicate some degree of consistency within the instrument. Because relatively low correlations were found between placement levels on cloze tests and GRI's there is reason to suggest that the two types of procedures do not measure the same skills. Both measures were, it appears, able to differentiate between frustration level readers and better readers whose scores fell within either an instructional or independent reading level. Peterson and Carroll (1974) reported similar findings when using the cloze test with disabled 6,7,8 and 9th grade subjects. The present study, using good readers enrolled in grades 7,8 and 9, supports their finding for the cloze test and also a group reading inventory which was not a part of their study.

The current research found social studies cloze test mean scores consistently differentiating between frustration level readers and instructional and/or independent level readers. Perhaps using the 37 percent - 57 percent range for the instructional level compared to a 38 percent - 44 percent used in several earlier studies accounts for some stability of mean scores.

Significant differences were found to exist between vocabulary and comprehension GMRT sub-test scores and GRI and cloze placement levels. Perhaps the widespread use of standardized achievement tests in reading does not provide a reliable source of data from which to draw information relative to content area reading abilities for junior high school readers. While this study dealt with social studies content, Peters, Peters and Kaufman (1975) report a similar conclusion in a study of reading comprehension in their

investigation involving four content areas.

The present study does not support or reject the use of Group Reading Inventories or cloze tests as diagnostic instruments for use in grades 7, 8 and 9 social studies classes. The data does in general, suggest that for a sample of good readers:

1. Properly constructed and administered cloze tests and group reading inventories in social studies can be used to accurately place readers in placement level categories. This is particularily true between frustration and instructional/independent levels. The division between instruction and independent level readers using these two tests was less clear.
2. Group reading inventories in social studies can be used to discriminate between varying degrees of reader ability to answer main idea, detail, context and inference type questions.
3. Cloze tests and group reading inventories do not necessarily measure the same reading skills.
4. Results based on comparisons using one popular standardized reading test indicate that standardized test scores differ significantly when compared with reading performance as measured by social studies cloze tests and group reading inventories.

5. If student textbook reading ability is important, classroom teachers should consider using informal types of reading assessment in addition to standardized reading achievement tests in grades 7,8 and 9.

REFERENCES

Allington, Richard. "Improving Content Area Instruction in the Middle School." Journal of Reading, 18:457, March, 1975.

Berger, H.I. "The Difficulty of Third Grade Readers." Elementary School Journal, 47:391-395, March 1947.

Bormuth, John. "The Cloze Readability Procedure." Elementary English, 45:429-436, April, 1968.

Barron, A. and Claybaugh, A.L. Using Reading to Teach Subject Matter. Charles E. Merrill Publishing Co., 1974.

Carvell, Robert L. "Constructing Your Own Textbook - Specific Study Skills Inventory." Reading World, 19:239-245, 1980.

Cox, J.M. "An Evaluation of the Reading Levels of Junior High School Students and the Readability Levels of Selected Modern Language Textbooks." Unpublished Doctoral Dissertation, Florida State University, 1970.

Dishner, Earnest K. and John E. Readance. "Getting Started: Using the Textbook Diagnostically," Reading World, 17:36-49, 1977.

Galloway, Percilla. "How Secondary Students and Teachers Read Textbooks." Journal of Reading, 17:216-219, December 1973.

Gates, Arthur I. and Walter MacGinitie, Gates-MacGinitie Reading Tests, 1965.

Hafner, Lawrence E. Developmental Reading in Middle and Secondary Schools, New York: Macmillan Publishing Company, 1977.

Kaiser, Robert. "The Group Diagnostic Inventory," Decision Making: An Imperative for Language Learning, Proceedings of the Seventh Annual Language Arts Conference, Memphis State University, January, 1975.

Koenke, Karl. "Evaluation by the High School Teacher." Journal of Reading, 16:220, 222-225, December, 1972.

Marksheffel, Ned O. Better Reading in the Secondary Schools. Ronald Press, 1966.

Miller, Wilma H. Reading Diagnosis Kit, New York: The Center for Applied Research in Education, 1974.

McWilliams, Lana and Rakes, Thomas. Content Inventories for English, Science and Social Studies, Kendall Hunt Publishers, 1979.

Peters, Charles W., Peters, Nathaniel A. and Kaufman, B. Darwin, "A Comparative Analysis of Reading Comprehension in Four Content Areas." Reading: Convention and Inquiry. 24th Yearbook of the National Reading Conference, 200-226, 1975.

Peterson, Joe and Carroll, Martha. "The Cloze Procedure as an Indicator of the Instructional Level for Disabled Students." Interaction: Research and Practice in College - Adult Reading. 23rd Yearbook of the National Reading Conference, 153-157, 1974.

Rakes, Thomas A. "A Comparative Readability Study of Materials Used to Teach Adults to Read," Adult Education, 2:192-202, Spring, 1973.

Rakes, Thomas and Lana McWillimas. "A Comparative Study of Informal Group Assessment Procedures and Standardized Test Performance for Adult Readers," Resources in Education, May, 1979, (ED163445).

Rakes, Thomas. "A Group Instructional Inventory," Journal of Reading, 8:595-598, May, 1975.

Rakes, Thomas and Lana McWilliams, "Bridging the Gap: Alternatives to Standardized Testing," English Journal, 67:46-50, October, 1978.

Ramsey, Ronald D. "An Analysis of the Readability and Difficulty of Instructional Materials in a Junior High School." Unpublished Doctoral Dissertation, University of Kansas, 1961.

Shepherd, David L. Comprehensive High School Reading Methods. Charles E. Merrill, 1973.

Smith, Edwin, Guice, Bill and Cheek, Martha. "Informal Reading Inventories for the Content Areas: Science and Mathematics." Elementary English, 49:659-666, May, 1972.

Strang, Ruth. Diagnostic Teaching of Reading. McGraw-Hill Co., 1964.

Vaughan, Joseph L., Jr. and Paula J. Gaus. "Secondary Reading Inventory: A Modest Proposal," Journal of Reading, 21:716-720, 1978.

Viox, Ruth G. Evaluating Reading and Study Skills in the Secondary Classroom. International Reading Association, 1968.

Wood, L.W. "Readability of Certain Textbooks." Elementary English, 31:214-216, 1954.

TABLE I

Correlations for Subjects in Grades Seven, Eight and Nine on GRI, Cloze and Gates-MacGinitie Tests*

	GRI Correct	MI Wrong	Detail Wrong	Contx Wrong	Infer Wrong	GRI Placement Level	GRI % Correct	Cloze % Correct	Cloze Placement Level	GM Speed	GM Accuracy	GM Vocab	GM Comp
GRI Correct	1.0	.7439	.8445	.5089	.8446	.9256	.9355	.6430	.6140	.4717	.4964	.5520	.5722
MI Wrong	.7439	1.0	.6020	.2019	.6236	.6763	.6529	.5867	.5325	.4466	.4491	.4564	.4216
Detail Wrong	.8445	.6020	1.0	.3262	.6461	.7626	.7459	.5826	.5482	.3832	.4128	.4863	.5098
Contx Wrong	.5089	.2019	.3262	1.0	.3070	.4041	.4498	.3843	.3428	.2478	.2788	.3198	.4265
Infer Wrong	.8446	.6236	.6461	.3070	1.0	.7690	.7719	.5559	.5154	.4498	.4657	.5050	.5520
GRI Placement Level	.9256	.6763	.7626	.4041	.7690	1.0	.9487	.5350	.5283	.4065	.4268	.4884	.5108
GRI % Correct	.9355	.6529	.7459	.4498	.7719	.9487	1.0	.5100	.5026	.3987	.4224	.4929	.5089
Cloze % Correct	.6430	.5867	.5826	.3843	.5559	.5350	.5100	1.0	.8934	.4253	.4708	.5733	.5337
Cloze Placement Level	.6140	.5325	.5482	.3428	.5154	.5283	.5026	.8934	1.0	.3831	.4103	.5340	.4802
GM Speed	.4717	.4466	.3832	.2478	.4498	.4065	.3987	.4253	.3831	1.0	.9119	.4854	.4939
GM Accuracy	.4964	.4491	.4128	.2788	.4657	.4268	.4224	.4708	.4103	.9119	1.0	.5582	.5635
GM Vocb	.5520	.4564	.4863	.3198	.5050	.4884	.4929	.5733	5340	.4854	.5582	1.0	.6887
GM Comp	.5772	.4216	.5098	.4265	.5520	.5108	.5089	.5337	.4802	.4939	.5635	.6887	1.0

*All values significant at .001 level

TABLE II

Correlation Between Cloze, GRI and Gates-MacGinitie Comprehension Sub-Test Performance for 7, 8 and 9th Graders*

	Cloze Placement	GRI Placement	GM Comp
Grade 7			
Cloze PL	1.0	.5116	.4488
GRI PL	.5116	1.0	.4139
Grade 8			
Cloze PL	1.0	.2412	.3369
GRI PL	.2412	1.0	.4520
Grade 9			
Cloze PL	1.0	.4028	.5979
CRI PL	.4028	1.0	.5360

*All comparisons significant at .001 level

TABLE III

ANOVA Comparing Functional Reading Level Placement
On A CLOZE Test and the GMRT in Grade Seven

GMRT	D.F.	F-Ratio	F-Prob
Speed	1	6.077	.0159
Accuracy	1	10.499	.0018*
Vocabulary	1	36.908	.0000**
Comprehension	1	19.423	.0000**

N= 79

*P < 1.0125
**P < .0000

TABLE IV

ANOVA Comparing Functional Reading Level
Placement On A GRI and the GMRT In Grade Seven

GMRT	D.F.	F-Ratio	F-Prob
Speed	2	1.719	.1861
Accuracy	2	2.684	.0748
Vocabulary	2	6.712	.0021*
Comprehension	2	7.976	.0007*

N=78

*P < .0125

TABLE V

ANOVA Comparing Functional Reading Level Placement
On A Cloze Test And The GMRT In Grade Eight

GMRT	D.F.	F-Ratio	F-Prob
Speed	2	2.697	.0739
Accuracy	2	2.806	.0667
Vocabulary	2	12.645	.0000*
Comprehension	2	7.048	.0016**

N=78

*P < .0001
**P < .0125

TABLE VI

ANOVA Comparing Functional Reading Level Placement
On A GRI and The GMRT In Grade Eight

GMRT	D.F.	F-Ratio	F-Prob
Speed	2	2.238	.1137
Accuracy	2	3.559	.0333
Vocabulary	2	8.999	.0003*
Comprehension	2	12.299	.0000**

N=79

*P < .0125
**P < .0001

TABLE VII

ANOVA Comparing Functional Reading Level Placement
On A Cloze Test And The GMRT In Grade Nine

GMRT	D.F.	F-Ratio	F-Prob
Speed	2	9.856	.0001*
Accuracy	2	12.277	.0000**
Vocabulary	2	17.717	.0000**
Comprehension	2	38.731	.0000**

N=121

*P < .0125
**P < .0001

TABLE VIII

ANOVA Comparing Functional Reading Level Placement
On A GRI and the GMRT In Grade Nine

GMRT	D.F.	F-Ratio	F-Prob
Speed	2	11.005	.0000*
Accuracy	2	11.603	.0000*
Vocabulary	2	15.384	.0000*
Comprehension	2	24.091	.0000*

N=121

*P < .0001